Teoría de las comunicaciones

Primer cuatrimestre de 2017

Departamento de Computación
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Trabajo práctico 1: Wiretapping

Integrante	LU	Correo electrónico
Daniel Rodriguez		danielca@gmail.com
Abel Delgadillo		adelgadillo91@gmail.com

Reservado para la cátedra

Instancia	Docente	Nota
Primera entrega		
Segunda entrega		

Comentarios del corrector:

Índice

Índice de figuras

1. Introducción

En este informe nos proponemos analizar diferentes redes locales utilizando herramientas derivadas de la teoria de la información. El objetivo es usar datos triviales de los paquetes que pasan por la red para descubrir patrones y topologias sin conocer previamente nada sobre los nodos de la red. Para lograr esto utilizamos direcciones de la capa de enlace en un principio y luego datos de paquetes `ARP` (address resolution protocol, un protocolo utilizado para obtener una dirección de enlace dada una direccion de red para que 2 nodos en una misma red local de acceso compartido puedan comunicarse) para obtener mediciones de información y entropia de la misma red modelada de modo diferente.

2. Métodos y ambientes

Las redes seran analizadas con dos herramientas que escuchan un medio compartido en modo promiscuo. La primer herramienta modela la red como una fuente S cuyos posibles simbolos son $S_{unicast}$ si el paquete leido es enviado a un nodo especifico y $S_{broadcast}$ si el paquete es enviado a la direccion de broadcast (`ff:ff:ff:ff:ff:ff`), ningún paquete es filtrado, cualquier paquete que llega nuestro enlace será tenido en cuenta.
La segunda herramienta modelará la red como una fuente $S1$ cuyos posibles simbolos son todas las direcciones de red disponibles en la red (como esto es desconocido, vamos a asumir que todos los simbolos observados son todos los simbolos posibles) en paquetes ARP que nos encontramos en la red. La consigna era definir una función que designe alguñ simbolo como distinguido fundado en algún resultado matematico y para esto distinguimos a aquellos simbolo o simbolos que tengan información menor a la entropía dado que estos seran más comunes de ver (por la definición de información y entropía), es decir, que los otros nodos de la red piden su dirección de enlace muy frecuentemente. Esto hará distinguir a nodos internos muy usados o bien al gateway por defecto de la red por la cual todos los nodos salen a otras redes como internet.

Las N redes que elegimos para experimentar son las siguientes

- Red hogareña: red local wi-fi en una casa particular, lo unico que sabemos es que contiene una cantidad muy limitada de computadoras personales y dispositivos mobiles.
- Red laboral: red local ethernet en una oficina, sabemos que conviven computadoras de empleados y servidores.
- Red pública: red wi-fi en un shopping, diferentes tipos de usuarios, dispositivos y duración de leasing the direcciones de red

3. Red hogareña

Esta red en principio es muy controlada y solo sabemos que cuenta con un número pequeño de nodos. La red usa en su mayoria wi-fi por lo cual es de esperar que el nodo que usamos para monitorear la red reciba amplia cantidad de paquetes who-has y is-at dado que no hay switches segmentando la red que filtren respuestas is-at de terceros.

El primer experimento, el cual ve la red como una fuente con 2 simbolos unicamente, arroja luego de monitorear la red durante 30 minutos que en ella la mayor parte del trafico es unicast y solo el 17 % es broadcast.

Figura 1: Información de S, Red hogareña

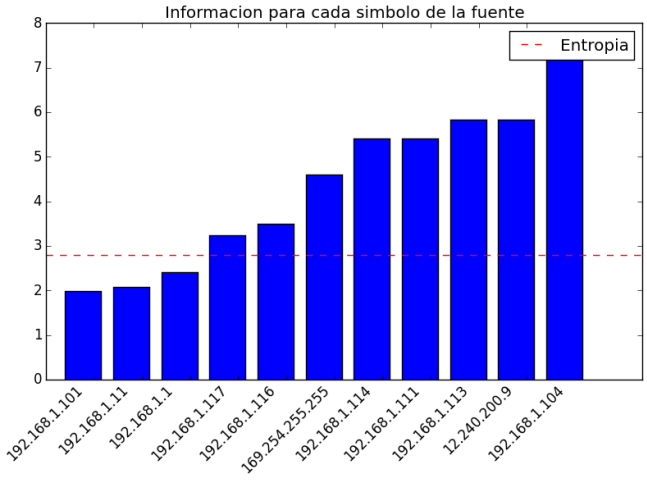
	frecuencia	información
$S_{broadcast}$	0.17	2.56
$S_{unicast}$	0.83	0.27

Estos valores nos dan una entropía de 0.66 bits (siendo el máximo 1 dado que hay 2 simbolos en principio equiprobables) lo cual concuerda con las expectativas dado que en una red local hogareña se espera ver mas que nada trafico unicast entre los nodos y el gateway hacia internet.

En el segundo experimento, el cual modela la red basado en la dirección a resolver en mensajes ARP, usamos la misma captura utilizada en el experimento anterior. El programa que usamos para analizar los resultados destaco 3 nodos de la red cuando esperabamos solo 1. El siguiente grafico muestra la información de cada simbolo y la entropía de la fuente, como se puede ver hay 3 nodos destacados bajo nuestra

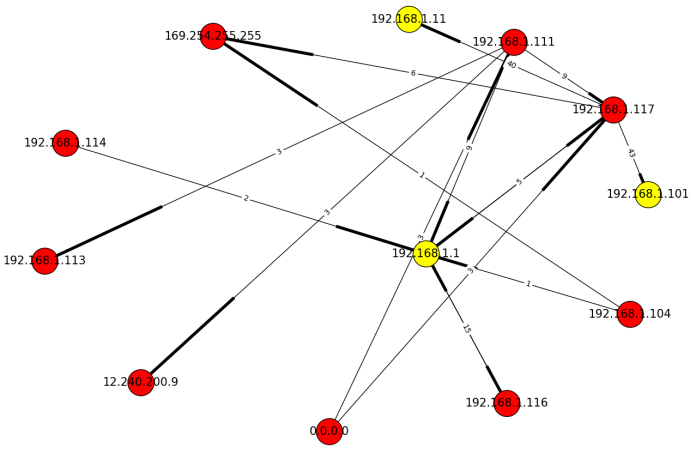
definición.

Figura 2: Información red hogareña



Una investigación más meticulosa concluyó que `192.168.1.1` es el gateway por defecto de la red mientras que los otros 2 nodos destacados, `192.168.1.11` y `192.168.1.101` ni siquiera existen en la red, las direcciones no fueron asignadas y ningún nodo las esta usando mas un nodo de la red esta constantemente intentando resolver su dirección de enlace. Solo podemos suponer que este host tiene software mal configurado o defectuoso. Las interacciones de los nodos se puede ver en más detalle con el siguiente grafico donde se puede ver a los nodos destacados y las relaciones con los otros nodos. Notar que en el grafico, asi como en el experimento también aparece un nodo con dirección `169.254.255.255` la cual es la dirección que el sistema operativo windows asignado a un nodo cuando este no puede contactar a ningún servidor DHCP para que le asigne una dirección libre. Usualmente los nodos operan con esta dirección durante segundos o minutos hasta que se puedan proveer de una dirección valida.

Figura 3: Visualización red hogareña, en amarillo los nodos destacados



El gráfico confirma que los nodos no existentes que se destacaron son solo accedidos por un nodo.

4. Red laboral

Para explorar nuevas topologias y relaciones decidimos monitorear una red laboral en una oficina de mediana envergadura. En esta red local conviven nodos de empleados, servidores de distinto uso, asi como impresoras de red y otros dispositivos de oficina.

El primer experimento, el cual ve la red como una fuente con 2 simbolos unicamente, monitoreo la red local a traves de un enlace ethernet durante 20 minutos. Como era de esperarse cerca del 100 % del trafico es unicast dado que es una red con mucho trafico hacia un servidor en particular o hacia el gateway.

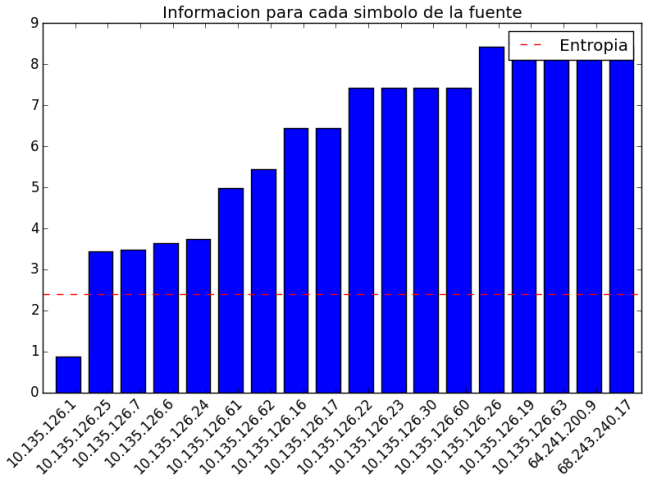
Figura 4: Información de S, Red laboral

	frecuencia	información
$S_{broadcast}$	0.03	5.06
$S_{unicast}$	0.97	0.04

Estos valores nos dan una entropía de 0.19 bits (siendo el máximo 1 dado que hay 2 simbolos en principio equiprobables), dado que el trafico broadcast sin ser despreciable solo es un porcentaje pequeño del trafico total.

En el segundo experimento, el cual modela la red basado en la dirección a resolver en mensajes ARP, usamos la misma captura utilizada en el experimento anterior. Bajo nuestra definición para destacar nodos (que la información del simbolo de ese nodo sea menor o igual a la entropía de la fuente) solo un nodo fue destacado por amplio margen, el gateway por defecto de la red.

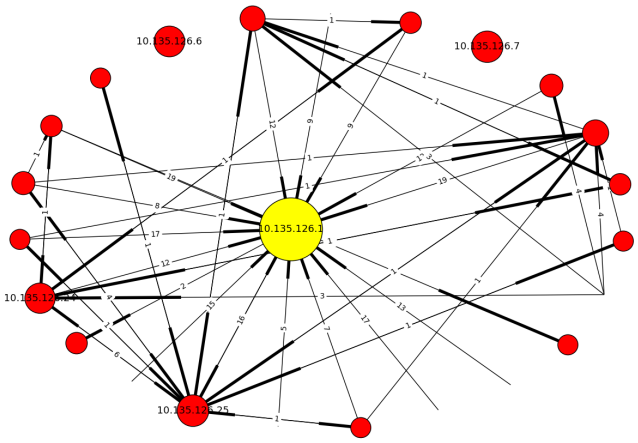
Figura 5: Información red laboral



La visualización de la red modelada como la fuente descriptiva muestra una topologia y relación entre nodos muy similar a la realidad en terminos de trafico total y disposición.

Un analisis manual del trafico mostró 2 fenomenos de ARP no usuales, ARP gratuitos y ARP de sondeo. Los ARP gratituitos pueden ser tanto `who-has` como `is-at` donde la dirección de origen como destino son las mismas (y la dirección de enlace destino es broadcast). El uso de los mismos es precargar o refrescar las tablas de otros nodos para evitar tener que traducir en tiempo real una dirección de red. Los ARP de sondeo tienen un fin similar, un nodo al que se le asigno manualmente o automaticamente

Figura 6: Visualización red laboral, en amarillo los nodos destacados



una dirección de red, envia un ARP con dicha dirección a la red y si algún nodo le responde entonces sabra que la dirección ya esta en uso y evitara la colisión de alguna forma. Ningún de estos tipos de mensajes estan oficialmente documentados pero forman parte de toda red que necesite autoregularse correctamente.

5. Red pública

Situados en un patio de comidas, monitoreamos la red wifi de un shopping importante durante 30 minutos con el objetivo de analizar una red con muy poco control y en constante cambio.

El primer experimento, el cual ve la red como una fuente con 2 simbolos unicamente, resulto en números similares a la red hogareña aunque en un principio se podria suponer que el trafico broadcast iba a ser mucho mas elevado con tantos nodos entrando y saliendo de la red.

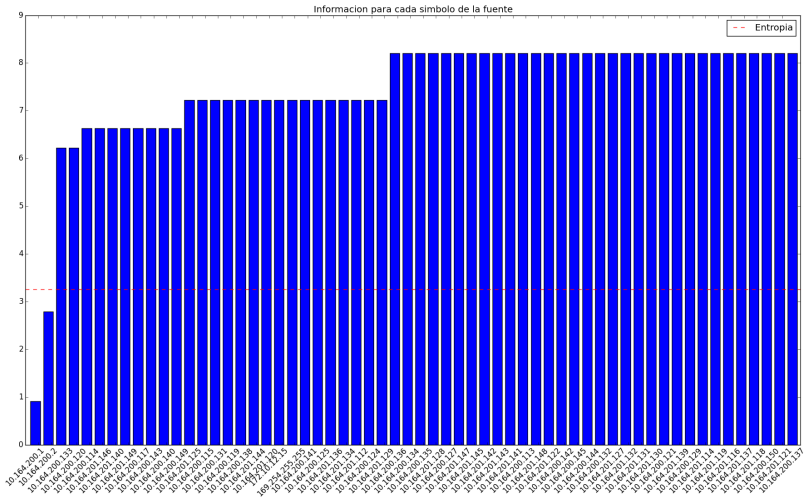
Figura 7: Información de S pública

	frecuencia	información
$S_{broadcast}$	0.21	2.25
$S_{unicast}$	0.79	0.34

Estos valores nos dan una entropía de 0.65 bits (siendo el máximo 1 dado que hay 2 simbolos en principio equiprobables). Suponemos que la similitud con la red hogareña esta en que el tiempo en que monitoreamos la red coincide o es menor que el tiempo de cada nodo en la red.

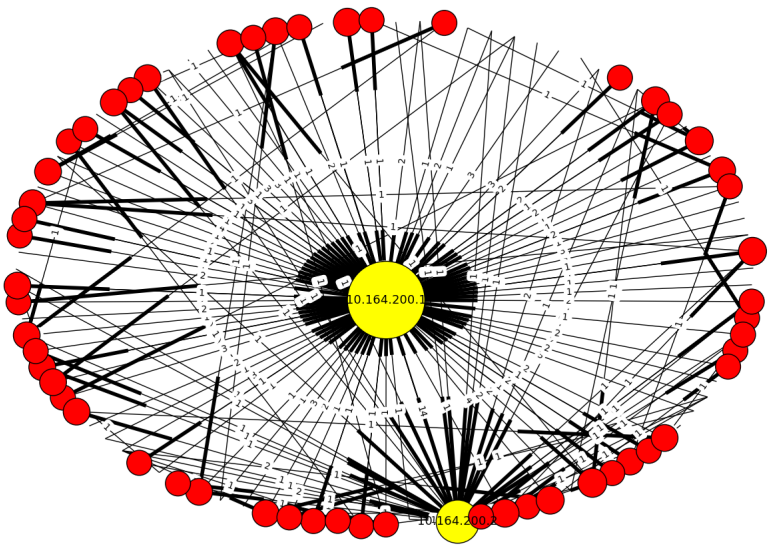
En el segundo experimento, el cual modela la red basado en la dirección a resolver en mensajes ARP, usamos la misma captura utilizada en el experimento anterior. Nuestra herramienta y definición de nodo destacado encontró 2 nodos para destacar, uno sabemos que es gateway por default mas no sabemos exactamente que es el segundo, posiblemente otro gateway para otro sector o diferentes clientes (dado que analizando la dirección de enlace podemos ver que es un router de reconocida marca)

Figura 8: Información red pública



Se pueden notar 3 grupos de nodos basados en la información: destacados, aquellos que fueron accedidos dentro de la red por otros nodos y una gran cola de nodos que solo figuran por un otro ARP capturado. Lo mismo se refleja en la visualización de la red a partir de la fuente S1 donde se puede ver como los nodos destacados son varias magnitudes más grandes que el resto.

Figura 9: Visualización red pública, en amarillo los nodos destacados



Al igual que la red laboral antes descripta volvimos a observar ARP gratuitos y ARP de sondeo provenientes del gateway y de algunos nodos.

6. Conclusiones

Luego de analizar 3 redes con configuraciones diferentes podemos concluir que el trafico unicast parece ser usualmente superior al broadcast si se monitorea una red durante el tiempo suficiente, tal vez por que las 3 redes tienen un proposito comun que es comunicar una red local al exterior u otra red. También podemos mencionar que nuestra elección de nodos distinguidos resultó ser muy útil, no solo para distinguir el usual gateway por default si no como vimos tambien para detectar otros gateways no asignados al nodo monitor o bien para diagnosticar nodos que no funcionan bien como vimos en el primer caso. Esta elección se valió exclusivamente en la entropía de la fuente la cual sirvió efectivamente como un limite al cual un simbolo/nodo debe acercarse para ser relevante en la red y que la entropiá parece ser mayor cuando mayor cantidad de nodos hay en la red, esto tiene sentido si se modela usando ARP dado que vamos a incrementar la cantidad de simbolos con poca frencuencia en nuestra fuente. Cabe destacar también el descubrimiento de tecnicas utilizadas por routers y nodos para comunicar datos a traves de ARP con el objetivo de mejorar la performance de la red, utilizando mensajes no requeridos por ningún otro nodo.